

**UNIVERSIDAD AUTÓNOMA DE MADRID  
ESCUELA POLITÉCNICA SUPERIOR**



**Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación**

**TRABAJO FIN DE GRADO**

**DETECCIÓN DE RITMO CARDÍACO MEDIANTE  
ANÁLISIS DE SECUENCIAS DE VÍDEO EN COLOR**

**Ana Martín Doncel.  
Tutor: José María Martínez Sánchez.**

**Junio 2018**



# DETECCIÓN DE RITMO CARDÍACO MEDIANTE ANÁLISIS DE SECUENCIAS DE VÍDEO EN COLOR

**Ana Martín Doncel**

**Tutor: José María Martínez Sánchez**



**Video Processing and Understanding Lab**

**Departamento de Tecnología Electrónica y de las Comunicaciones**

**Escuela Politécnica Superior**

**Universidad Autónoma de Madrid**

**Junio 2018**

Trabajo parcialmente financiado por el Ministerio de Economía y Competitividad del Gobierno de España bajo el proyecto TEC2014-53176-R (HAVideo) (2015-2017)







# Resumen

El objetivo de este trabajo de fin de grado es implementar un sistema que estime, mediante el análisis de variaciones de color en la piel, la frecuencia cardíaca de un individuo, mediante técnicas no invasivas. Este estudio se ha visto motivado por la importancia que tiene el corazón en nuestra salud, al ser el encargado de bombear sangre a todos los rincones de nuestro organismo.

El tratamiento de imágenes, en cuanto a la extracción de información, está al orden del día en todas las aplicaciones médicas, debido al auge tecnológico actualmente presente. Es por esto, que el análisis de secuencias de vídeo, resulta una buena iniciativa para detectar el ritmo cardíaco de manera no invasiva.

La circulación de la sangre causa cambios volumétricos en los vasos sanguíneos, lo que produce que la trayectoria de la luz varíe de manera periódica durante el ciclo cardíaco. Además, la hemoglobina posee un rango espectral visible fijo, por lo que resulta una buena idea trabajar con imágenes de color para la extracción de la frecuencia cardíaca.

Este trabajo se ha dividido en dos partes. En la primera parte se han diseñado varios algoritmos que siguen el fundamento anterior como base y se han comparado con un algoritmo ya existente que fue previamente probado y demostrado, para observar la fiabilidad que poseen. Se obtiene un rango de porcentajes de error de entre 0.339 y 3.543, entre los algoritmos.

Una vez realizados los algoritmos, y habiendo comprobado buenos resultados, se procede a realizar la segunda parte del trabajo, que consiste en evaluar los distintos parámetros variables que los contienen, mediante un barrido. Mediante la aplicación de este estudio se han llegado a unos resultados satisfactorios.

# Palabras clave

Ritmo cardíaco, región de interés, método Eurelian, espacio transformado, cambios de color, detección, análisis de secuencias de vídeo.



# Abstract

The aim of this Bachelor Thesis is to develop a system which estimates heart rate from a random person analysing variations un skin colour by using non-invasive techniques. This study has been motivated by the importance of our heart's health, since it is in charge of pumping blood to every single part of the body.

Image processing, in particular information extraction, is the order of the day in medical applications, due to technological advances nowadays. For this reason, video sequence analysis seems to be a good initiative to detect heart rhythm in a non-invasive way.

Blood circulation causes volumetric changes in blood vessels. This makes the light path vary in a periodic way throughout the cardiac cycle. Furthermore, haemoglobin has a fixed and visible spectral range so it would be a good idea to work with colorful images to extract heart rate.

This thesis has been divided into two parts. In the first one we have designed algorithms which are based on the method that has been described. They have been compared to an already existing algorithm which has been tested so we can observe their reliability, reaching a percentage rate from 0.339 y 3.543.

Once we have completed this experiments and results have been obtained, we proceed to the second part. This part consists in an evaluation of different parameters through a sweep of the signal. In this study we have achieved good results.

# Keywords

Heart rate, area of interest, Eurlen method, area transformed, colour changes, detection, video sequence analysis



# Agradecimientos

*Agradecer a todas aquellas personas que me han acompañado durante estos años de carrera. A mis padres y mi pequeña, por estar siempre ahí, apoyarme de manera incondicional y confiar en mi potencial.*

*Agradecer también a mi tutor, Chema, por haberme dado la oportunidad de realizar este trabajo y de haberme proporcionado su ayuda y apoyo. Doy las gracias también al VPU y a la gente que ha estado conmigo ahí.*

*A mis compañeros de carrera y de vida, que han compartido conmigo momentos de agobio, estrés y lloros encerrados entre estas cuatro paredes que componen la escuela politécnica superior; en especial a Claudia y Raúl con los que he compartido los peores y mejores momentos de la carrera.*

*A Sergio, por ayudarme siempre en mi camino, ser mi compañero en general y un chofer espectacular.*

*Un especial agradecimiento a Pato, la cual, aún sin ser una persona muy empática ha sabido como llegar a animarme para conseguir esto, en estas semanas que han sido, para mí, de las más duras. Gracias.*

*Ya sabéis los que me conocéis que soy una persona muy sensible, y si tuviera que nombrar a todos aquí no acabaría. Os llevo muy dentro, y doy gracias por haberme encontrado en mi camino con todos vosotros.*



# Índice general

<b>Resumen</b>	<b>V</b>
<b>Abstract</b>	<b>VII</b>
<b>Agradecimientos</b>	<b>IX</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivos . . . . .	2
1.3. Organización de la memoria . . . . .	2
<b>2. Estado del arte</b>	<b>5</b>
2.1. Introducción . . . . .	5
2.2. Anatomía del corazón . . . . .	6
2.3. Métodos de obtención del ritmo cardíaco mediante análisis de vídeo . . . . .	8
2.3.1. Obtención de ritmo cardíaco mediante cambios de color . . . . .	9
2.3.2. Obtención de ritmo cardíaco mediante movimiento . . . . .	12
2.4. Kinect y WebCam . . . . .	13
2.5. Conclusiones . . . . .	14
<b>3. Diseño y desarrollo</b>	<b>15</b>
3.1. Introducción . . . . .	15
3.2. Algoritmo color . . . . .	16
3.2.1. Diseño . . . . .	16
3.2.2. Desarrollo . . . . .	17
3.3. Algoritmo Luminancia . . . . .	22
3.3.1. Diseño . . . . .	22
3.4. Desarrollos no incluidos . . . . .	23
3.4.1. Introducción . . . . .	23
3.5. Validación de resultados . . . . .	23
3.5.1. Introducción . . . . .	23
3.5.2. Resultados . . . . .	24
3.5.3. Conclusiones . . . . .	24

<b>4. Evaluación</b>	<b>27</b>
4.1. Introducción . . . . .	27
4.2. Marco de evaluación . . . . .	27
4.2.1. <i>Ground truth</i> . . . . .	27
4.2.2. <i>Dataset</i> . . . . .	28
4.2.3. Métricas . . . . .	29
4.3. Pruebas y resultados . . . . .	30
4.3.1. Pruebas y resultados barrido . . . . .	30
4.3.2. Resultados finales . . . . .	37
4.4. Conclusión . . . . .	37
<b>5. Conclusiones y trabajo futuro</b>	<b>39</b>
5.1. Conclusiones . . . . .	39
5.2. Trabajo futuro . . . . .	39



# Índice de figuras

2.1. Circulación de la sangre desde el corazón hacia la cabeza mediante las arterias carótidas. . . . .	6
2.2. Ciclo cardíaco. . . . .	7
2.3. Intervalo RR. . . . .	7
2.4. Electrocardiograma de una persona con Arritmia . . . . .	8
2.5. Electrocardiograma de una persona con Taquicardia . . . . .	8
2.6. Electrocardiograma de una persona con Bradicardia . . . . .	8
2.7. Cardio App . . . . .	9
2.8. Esquema base: Detección del ritmo cardíaco mediante variaciones de color . . . . .	10
2.9. Algoritmo basado en variaciones de color con separación RGB y uso de ICA. . . . .	12
2.10. Balistocardiograma. . . . .	13
2.11. Kinect. . . . .	14
3.1. Etapas del algoritmo de color . . . . .	16
3.2. Zonas de Interés . . . . .	16
3.3. Esquema de amplificación de vídeo - Fuente [ref MIT] . . . . .	17
3.4. Modelo de bloques de la variante 1 . . . . .	18
3.5. Modelo de bloques de la variante 2 . . . . .	18
3.6. Cara Detectada . . . . .	19
3.7. Representación piramidal del análisis multiresolución. Fuente [6] . . .	20
3.8. Amplificación de variaciones de color . . . . .	21
3.9. Representación de la fft para la extracción del pulso cardíaco . . . . .	22
3.10. Imágenes de prueba . . . . .	24
4.1. Smartwatch “Fitbit Versa” . . . . .	28
4.2. Muestras de los distintos tipos de secuencias que componen el dataset	29
4.3. Esquema del algoritmo que se llevará a cabo para realizar el barrido .	31
4.4. Canal rojo - Banda frecuencia 0.75-0.833 Hz . . . . .	32
4.5. Canal rojo - Banda frecuencia 1-1.0833 Hz . . . . .	33
4.6. Canal rojo - Banda frecuencia 1.1667-1.25 Hz . . . . .	34
4.7. Gráficas de resultados . . . . .	36
4.8. Posibles valores de pulsaciones . . . . .	36



# Índice de tablas

3.1. Validación de resultados con dataset MIT . . . . .	24
4.1. Ground Truth . . . . .	29
4.2. Parámetros Seleccionados . . . . .	35
4.3. Resultados finales con utilización DCT y FFT . . . . .	37



# Capítulo 1

## Introducción

### 1.1. Motivación

El corazón es uno de los músculos más importantes de nuestro cuerpo, ya que, es el encargado de bombear sangre a todos los órganos. Por este motivo, la frecuencia con la que bombea es uno de los signos vitales más interesantes de estudiar y analizar en el diagnóstico clínico.

Existe un interés creciente en la extracción de la frecuencia cardíaca sin la necesidad de contacto físico, sobre todo, en el caso particular de recién nacidos, ancianos o personas con alguna problemática sobre la piel, haciéndola más frágil o sensible a los métodos convencionales, además de ser útil para personas que presentan haptefobia<sup>1</sup>. Para todas estas casuísticas es muy interesante el estudio de técnicas no invasivas como es el caso de la detección del ritmo cardíaco mediante el análisis de secuencias de vídeo.

Como se observará en el siguiente capítulo, existen diferentes técnicas de obtención del pulso cardíaco mediante tratamiento de vídeo que se basan en el estudio de las sutiles oscilaciones de la cabeza -tratamiento de vídeo mediante movimiento- y sutiles cambios de color de la piel -tratamiento de vídeo mediante diferencias de color-, imperceptibles para el ojo humano, que se producen a causa del flujo sanguíneo.

El fin de este trabajo es implementar y analizar diversos algoritmos que nos puedan permitir detectar el ritmo cardíaco mediante el análisis de imágenes en color. Se busca diseñar un algoritmo que trabaje sobre las diferencias de color que se producen en la piel, y un segundo algoritmo que trabaje basándose en el movimiento de la cabeza debido a la aceleración de la sangre a través de la aorta.

La motivación de este trabajo, además de estimar el pulso cardíaco sin necesidad de

---

<sup>1</sup>Personas que sufren fobia al contacto físico.

usar técnicas invasivas, es poder estimarlo con el menor retardo posible, ya que a medida que la población envejece, la monitorización continua, o al menos frecuente fuera de los entornos clínicos, puede proporcionar a los médicos no solo muestras oportunas sino también tendencias a largo plazo y análisis estadísticos.

## 1.2. Objetivos

El objetivo principal de este trabajo consiste en crear un sistema que estime la frecuencia cardíaca de las personas, en nuestro caso, mediante el análisis de secuencias de vídeo en color. Este algoritmo abordará el análisis del color de piel y seguimiento de zonas de interés a lo largo del tiempo. Se analizará en detalle la robustez en función de la calidad del vídeo, además de analizar la información que aportan las distintas regiones de interés de la cara, y la combinación de estas para llegar a buenos resultados.

A continuación, se explicarán brevemente los pasos que contribuirán de manera progresiva a la obtención del objetivo total:

- Estudio del estado del arte.
- Desarrollo de un conjunto de datos de evaluación (*dataset*).
- Desarrollo de una metodología de evaluación de resultados.
- Implementación de un algoritmo base de análisis de color.
- Modificación de este para luminancia.
- Implementación de algoritmo de análisis de movimiento en imágenes de color.
- Evaluación de la viabilidad del desarrollo de un sistema combinado en función de la distancia de la resolución de la cámara.

## 1.3. Organización de la memoria

La memoria consta de los siguientes capítulos:

- Capítulo 1. Motivación, objetivos y organización de la memoria.
- Capítulo 2. Estado del arte en la obtención del ritmo cardíaco mediante el análisis de vídeo.
- Capítulo 3. Diseño y desarrollo de los distintos algoritmos realizados durante este trabajo. Evaluación de robustez de dichos algoritmos.

- Capítulo 4. Descripción del *dataset* generado. Descripción de las pruebas realizadas y discusión sobre los resultados obtenidos.
- Capítulo 5. Conclusiones y trabajo futuro.
- Bibliografía.





## Capítulo 2

# Estado del arte

### 2.1. Introducción

Para la correcta realización de este trabajo de fin de grado se ha estudiado con detenimiento distinta bibliografía relacionada con la detección del ritmo cardíaco.

El estudio del ritmo cardíaco se remonta a decenas de siglos atrás. Desde tiempos anteriores a Cristo, se encuentran numerosas referencias [8] a los pulsos del corazón en distintas civilizaciones como fueron la india, china, egipcia o griega. En estas civilizaciones ya se describían variedades de impulsos durante diferentes procesos fisiológicos y patológicos. Además, se tiene constancia de una ciencia médica que describía los distintos rangos de frecuencia cardíaca según los distintos grupos de edad[8].

No fue hasta el siglo XVI cuando W.Harvey [9] describió por primera vez que el flujo de la sangre seguía un recorrido circular por nuestro cuerpo; desde el corazón hasta las extremidades a través de las arterias retornando o al corazón mediante las venas. En sus estudios hace referencia a las fases de contracción y relajación del corazón, principio fundamental en la medicina moderna para la obtención de la frecuencia cardíaca.

El principio básico de la detección de este signo vital mediante análisis de vídeo es el mismo; debido al flujo sanguíneo y las distintas fases del ciclo cardíaco se producen sutiles oscilaciones y cambios de color en nuestro cuerpo, que son más acentuados en la cabeza debido a nuestra anatomía, mediante los cuales podemos extraer información acerca de la actividad cardíaca.

Además de proporcionar una vía no intrusiva de calcular el pulso cardíaco, este método puede ser utilizado para extraer información acerca de la actividad cardíaca de un individuo concreto, como los sutiles cambios en la duración de los latidos

del corazón que están asociados con la salud del sistema nervioso[1]. Al tratarse de un pilar básico en este trabajo, en la siguiente sección se explicará de manera más detallada el funcionamiento del ciclo cardíaco.

Estos cambios no son perceptibles para el ojo humano, pero gracias al gran desarrollo del procesamiento de vídeo se pueden percibir y obtener de ellos la información necesaria para calcular el ritmo cardíaco. Durante la realización de este estudio se ha observado que existen diferentes métodos de obtención del ritmo cardíaco mediante el análisis del vídeo, pudiéndose agrupar estos en dos colectivos: obtención en base a las variaciones de color y en base al movimiento.

## 2.2. Anatomía del corazón

El corazón es un músculo hueco que bombea sangre por todo el cuerpo con contracciones cíclicas y uniformes. Está compuesto por cuatro cámaras: el ventrículo izquierdo / derecho y la aurícula izquierda / derecha 2.1

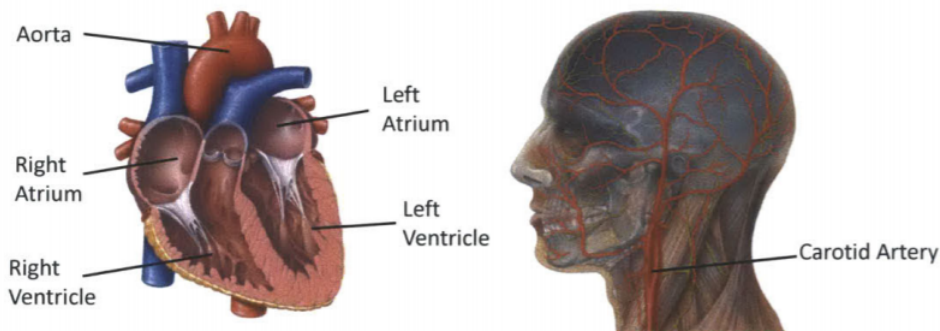


Figura 2.1: Circulación de la sangre desde el corazón hacia la cabeza mediante las arterias carótidas situadas en ambos lados de la cabeza.

Durante la primera fase del ciclo cardíaco conocido como diástole, los ventrículos se relajan y permiten que la sangre fluya hacia ellos desde las aurículas. En la siguiente fase conocida como sístole, los ventrículos se contraen y bombean sangre a las arterias pulmonares y aorta. La aorta a su vez transporta sangre al resto del cuerpo. La cabeza y el cuello reciben sangre de la aorta a través de las arterias carótidas comunes, que además se dividen en las arterias carótidas internas y externas en el cuello. El movimiento de la cabeza y los cambios de color en la piel en esta, se atribuyen a la gran aceleración de la sangre en la aorta y las carótidas[3].

La duración del ciclo cardíaco de un individuo se puede definir por el tiempo transcurrido desde un punto del ciclo cardíaco hasta el mismo punto del siguiente ciclo, es

decir, desde el inicio de la diástole hasta el inicio de la siguiente diástole tal y como se puede observar en 2.2.

Por este motivo la frecuencia cardíaca se puede medir como la cantidad de ciclos que transcurren en un periodo de tiempo seleccionado. Se usa principalmente para detectar problemas agudos.

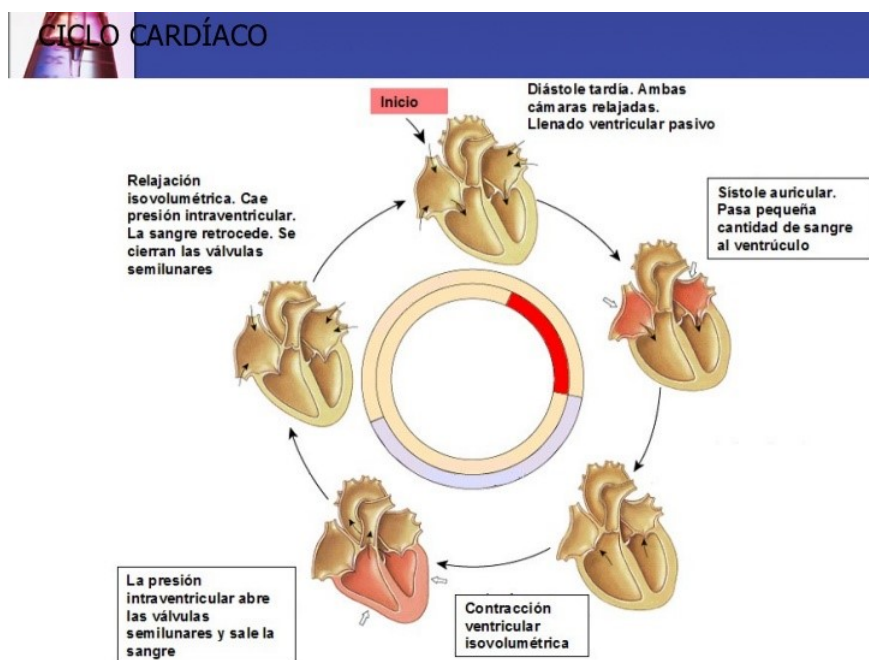


Figura 2.2: Ciclo cardíaco.

Como se puede observar en 2.3, la frecuencia cardíaca analizada en un electrocardiograma se puede medir como la distancia entre dos ondas “R” sucesivas, también denominado intervalo RR. De manera sencilla, se puede entender una onda R como la onda representada en el electrocardiograma con mayor amplitud. Por cada latido del corazón tendremos una onda R, y la distancia entre un latido y otro consecutivo será similar si el ritmo es regular.

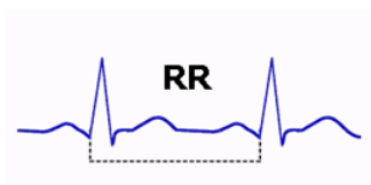


Figura 2.3: Intervalo RR.

Si el ritmo es irregular, como es el caso de personas que sufren arritmias, la

distancia entre estas ondas no será constante (2.4). Además, existen afecciones, como son las taquicardias y bradicardias (2.5,2.6), en las cuales los intervalos RR son de duración mayor o menor a lo establecido como ritmo sinusal normal (entre 60 y 100 latidos por minutos), esto implica que los latidos del corazón están acelerados o retardados.

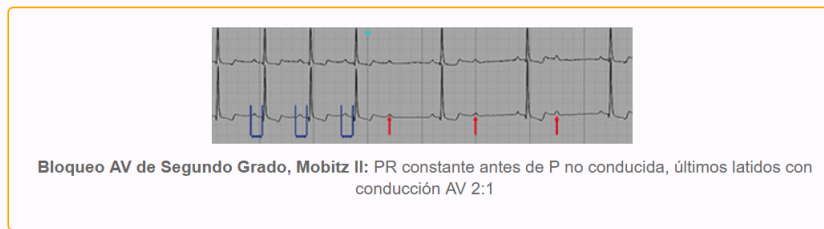


Figura 2.4: Electrocardiograma de una persona con Arritmia

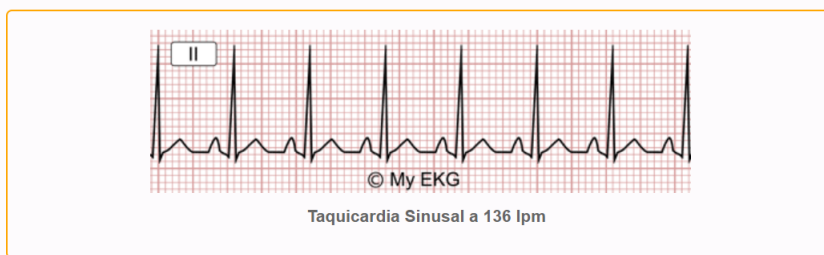


Figura 2.5: Electrocardiograma de una persona con Taquicardia

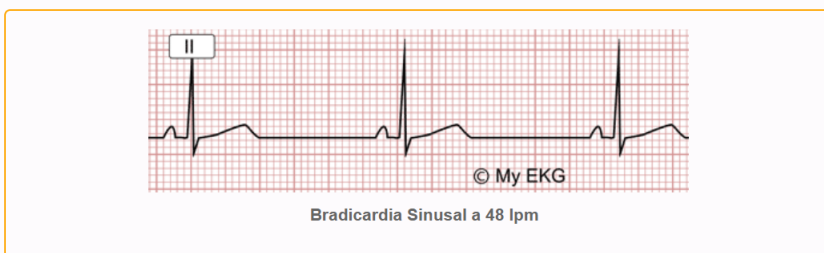


Figura 2.6: Electrocardiograma de una persona con Bradicardia

### 2.3. Métodos de obtención del ritmo cardíaco mediante análisis de vídeo

Las técnicas de detección del ritmo cardíaco mediante análisis de vídeo, tanto las de movimiento como las basadas en color, utilizan el mismo principio que se utiliza

en el análisis de los latidos del corazón mediante electrocardiograma explicado en el apartado anterior. La observación y análisis de las ondas R, nos aporta información sobre el ritmo cardíaco de las personas. Para llegar a una señal sencilla de tratar y de la que se pueda extraer esta información, en función de los parámetros observados (color, movimiento), llevan a cabo distintos procesos que se explicarán en este apartado. Estas técnicas se llevan usando ya desde hace más de una década, pues fue en 2002[15], cuando se empezaron a utilizar aplicándolas a secuencias de vídeo obtenidas mediante cámaras térmicas. En 2007 [21, 7] ya se realizaban estudios sobre la detección del ritmo cardíaco mediante secuencias de vídeo en color con cámaras convencionales, pero no fue hasta la primera publicación del Instituto Técnico de Massachusetts (MIT) [27], sobre el comportamiento de la piel y sus variaciones de color producidas por el flujo cardíaco, cuando empezó el crecimiento de esta técnica, momento en el cuál se empezaron a descubrir los movimientos que generaba nuestro cuerpo por esta misma razón.

### 2.3.1. Obtención de ritmo cardíaco mediante cambios de color

Este método es uno de los más usados en la actualidad, existen gran variedad de aplicaciones en el mercado que detectan el pulso cardíaco basándose en las variaciones de color, imperceptibles para el ojo humano, que se producen en nuestro cuerpo debido a la circulación sanguínea. Encontramos aplicaciones como la de Cardiio App 2.7 que posee dos modos: modo facial, y modo huella. El modo huella funciona situando el dedo sobre el sensor de la cámara y el modo facial situándose en frente de la cámara inmóvil mientras la aplicación extrae el pulso cardíaco.



Figura 2.7: Cardiio App.

La circulación sanguínea causa cambios volumétricos en los vasos sanguíneos que modifican la trayectoria de la luz. Esta es la premisa básica de la pletismografía, técnica basada en la medición de cambios de presión y volumen, que se utiliza para medir parámetros orientados al diagnóstico de enfermedades pulmonares y cardiovasculares. Los sensores rojos, verdes y azules (RGB) de una cámara de vídeo pueden captar esta señal pletismográfica mezclada con otras variaciones de luz causadas por el resto de los objetos que se encuentran en la escena.

Además, debido a que el color de la hemoglobina varía en torno a un rango espectral visible de la luz[10], cada uno de estos sensores registra una mezcla de estas fuentes en diferentes proporciones. La tarea de extraer el pulso, se basa en reconstruir la señal de interés de estos canales una vez se hayan amplificado de manera adecuada, para así obtener las variaciones de color que se producen en la piel durante las diferentes fases del ciclo cardíaco. Para obtener esta frecuencia cardíaca la gran mayoría de las técnicas basadas en color buscan la máxima energía en un espacio transformado, para encontrar las ondas R que se producen en cada latido del corazón, explicadas en el apartado anterior.

Durante el estudio del estado del arte se han encontrado más de 15 referencias relativas al análisis de secuencias de vídeo en color, que se dispondrán en la bibliografía de este trabajo. Todos ellos se basan, explicado de manera simplificada, en:

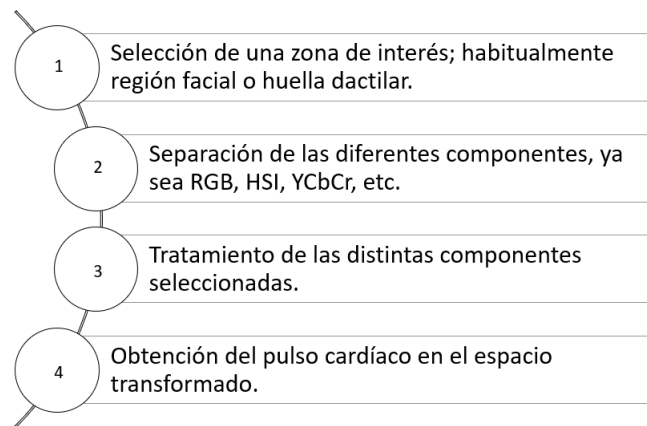


Figura 2.8: Esquema base: Detección del ritmo cardíaco mediante variaciones de color

Aunque todas las técnicas sigan este esquema, dentro de cada paso existen distintas variantes, dando así, una amplia gama de posibles métodos para la extracción de la frecuencia cardíaca mediante análisis de secuencias de vídeo en color, como pueden ser:

- Comparación de frames en el espacio de color HSI (Hue, Saturation, Intensity)[12].
- Separación en componentes de color RGB (Red, Green, Blue), para posterior procesamiento aplicando Análisis de componentes Independientes (ICA) o Análisis de Componentes Principales (PCA)[18, 11, 13, 14, 17, 20, 19, 16].
- Separación en componente de color RGB para posterior procesamiento sin aplicar ICA o PCA[26, 24, 22].
- Obtención de la intensidad o uso de la banda G del rostro para después aplicar análisis espectral auto regresivo[4].
- Diferencia de intensidad de brillo [21].
- Separación de las distintas componentes de color RGB o YCbCr y uso de descomposición piramidal[27].

Este último método es uno de los tratamientos más interesantes sobre secuencias de vídeo. Fue el publicado por el Instituto Técnico de Massachusetts (MIT), en 2012, y consiste en el aumento de vídeo, usando el método Eurelian, con la intención de revelar cambios imperceptibles para el ojo humano. Su objetivo era revelar variaciones temporales en vídeos que son difíciles o imposibles de revelar con el ojo humano. Tomaban como entrada una secuencia de vídeo, le aplicaban una descomposición espacial y, a continuación, filtraban la secuencia temporalmente. La señal resultante se amplificaba desvelando información que a priori era imperceptible para el ojo humano. Usando este método, eran capaces de observar el flujo sanguíneo a medida que circulaba por la cara. Esto tiene un gran potencial para aplicaciones de monitorización y diagnóstico clínico, donde, por ejemplo, la asimetría en el flujo sanguíneo puede implicar síntomas de problemas arteriales.

Además, uno de los trabajos que ha ganado mayor interés en los últimos años, consiste en separar la señal pletismográfica del ruido presente en la escena grabada mediante el análisis de las componentes independientes (ICA) nombradas anteriormente. Este método 2.9 promedia espacialmente las componentes RGB del área facial en cada *frame* del vídeo para formar tres señales, y se utiliza el análisis ICA para descomponer las señales en tres fuentes independientes. La fuente con el pico más alto en el espectro de potencia se elige como la señal que nos dará el pulso. Finalmente, la señal se suaviza y se interpola a 256 Hz para el análisis VFC (Variabilidad de la frecuencia cardíaca)[1].

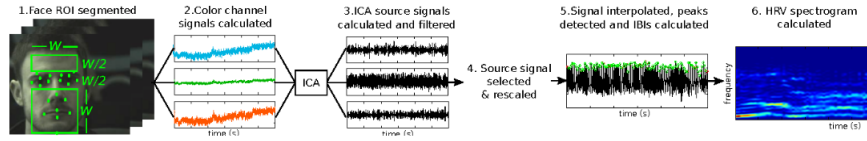


Figura 2.9: Algoritmo basado en variaciones de color con separación de las distintas componentes de color y uso de ICA.

También se observan distintos espacios transformados a partir de los cuales calcular la frecuencia cardíaca, entre los cuales dominan, el cálculo de la Transformada Discreta de Fourier, la Densidad Espectral de Potencia y la Transformada Discreta del Coseno.

Como limitación tendríamos que todos estos métodos requieren que la zona a analizar (zona facial, zona dactilar) sea visible para la cámara, y que la iluminación sea lo suficientemente buena como para detectar las componentes de color de manera adecuada.

### 2.3.2. Obtención de ritmo cardíaco mediante movimiento

Este método es una alternativa a la extracción del pulso cardíaco mediante análisis de secuencias de vídeo en color. El principio básico consiste en el movimiento cíclico de la sangre desde el corazón hacia la cabeza por la arteria abdominal y las arterias carótidas, que produce un movimiento periódico de la cabeza. Este movimiento es causado por la aceleración de la sangre por estas arterias. Si bien, es cierto, que estos movimientos son imperceptibles para el ojo humano, si amplificamos el vídeo se podrá observar este movimiento periódico.

La idea de explotar el movimiento inducido por el flujo sanguíneo no es nueva. En 1877, Gordon, registró los movimientos tanto horizontales como verticales que se producían en pacientes en una cama suspendida por cuatro cuerdas, este movimiento, se originaba debido a la energía transmitida desde el corazón hacia las grandes arterias, como consecuencia de la salida de la sangre desde los ventrículos y del retroceso del cuerpo, como reacción.

Pero no fue hasta 1939 cuando se elaboró el primer balistocardiografo<sup>2.10</sup>, que registraba el movimiento y además lo imprimía sobre papel fotográfico.



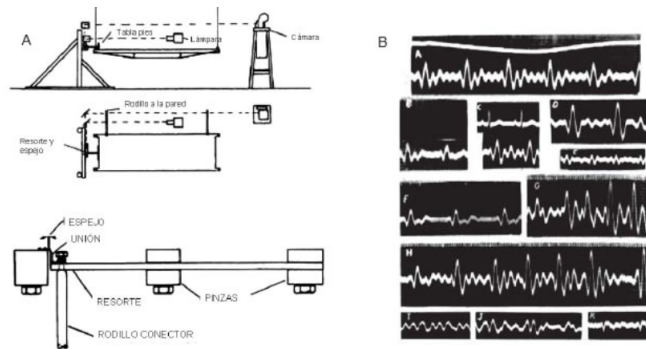


Figura 2.10: Balistocardiograma. Fuente [meter ref]

Hace unos años, He et al.[5] propusieron explotar este movimiento, mediante un dispositivo basado en acelerómetros, para monitorizar la actividad cardíaca.

El principio base de los algoritmos de movimiento consiste en encontrar la zona de interés (cabeza de la persona), rastrear los puntos característicos durante toda la secuencia del vídeo, filtrar sus velocidades en una banda de frecuencia temporal de interés, y usar el análisis de componentes principales (PCA) para encontrar una señal periódica causada por el pulso. A continuación, se extrae una frecuencia de pulso promedio de la señal al examinar su espectro y se obtienen la ubicación de latido precisas con un algoritmo de detección de pico simple.

El primer trabajo que se realizó de manera satisfactoria acerca de la detección del ritmo cardíaco sobre secuencias de vídeo mediante análisis de movimiento fue el publicado en 2013 por el Instituto Tecnológico de Massachusetts (MIT)[2]. En este estudio el individuo permanece inmóvil frente a la cámara mientras es grabado [2, 23, 1].

A partir de ahí, solo se producen modificaciones del algoritmo propuesto por el MIT para conseguir mejoras, se plantean soluciones para situaciones en las que la persona realiza movimientos de manera voluntaria para simular situaciones en las cuales el sujeto no permanece inmóvil, y posibles mejoras para trabajar con la Kinect v1 y v2, tanto en imágenes en profundidad como para imágenes en color .

## 2.4. Kinect y WebCam

En este trabajo se usarán secuencias en color obtenidas tanto de la Kinect v2 como de la cámara web de un ordenador.

La Kinect, diseñada y comercializada por Microsoft en el año 2010, fue concebida como un controlador de videojuegos para la videoconsola Xbox fabricada por esta misma empresa. Actualmente existen dos versiones de este producto; Microsoft

Kinect v1 y Microsoft Kinect v2. La Kinect consiste en un conjunto de cámaras, colocadas en una barra horizontal de 23 cm, con la capacidad de capturar imágenes en color, infrarrojo y en profundidad.



Figura 2.11: Kinect.

## 2.5. Conclusiones

Por todo lo comentado con anterioridad, se ha decidido trabajar con imágenes de color, ya que permite estudiar y analizar ambos métodos.

Por una parte, podremos trabajar sobre el método de detección del ritmo cardíaco mediante análisis de color, y de este modo, se podrá, en un futuro, no solo extraer el pulso cardíaco sino analizar los cambios de color y la distribución en la cara para detectar posibles síntomas de enfermedades cardíacas y arteriales, y por otro parte, además, podremos trabajar sobre el estudio de detección del ritmo cardíaco mediante análisis de movimiento sobre imágenes en color para mejorar la toma de decisiones en función de la distancia, luminancia y calidad del vídeo.

## Capítulo 3

# Diseño y desarrollo

### 3.1. Introducción

Tras estudiar los distintos algoritmos expuestos en el apartado anterior, describiremos de manera detallada el diseño y desarrollo de los algoritmos realizados en el presente trabajo.

En primera instancia se ha desarrollado un algoritmo basado en un estudio del MIT [23] consistente en la amplificación de video usando método Eulerian con la intención de revelar cambios imperceptibles para el ojo humano. Como entrada se han utilizados diversos videos en color de personas con diferentes tonalidades de piel y una vez amplificados los cambios de color en la cara del usuario se ha realizado un análisis sobre las componentes de color RGB (Red, green, blue) para extraer el ritmo cardiaco de la persona.

El segundo algoritmo desarrollado tiene una estructura muy similar, la diferencia principal es que el tratamiento de las imágenes se realizará utilizando, en lugar de las componentes RGB, las componentes de color YCbCr (Luminancia y crominancias). Este es un espacio de color muy utilizado en las técnicas de fotografía y vídeo, ya que, permite realizar tratamientos sobre imágenes con un tratamiento y codificación más eficientes.

Por último, se ha implementado un código de detección del pulso cardíaco mediante movimiento para imágenes en color. Dicho algoritmo es una mejora del presentado por Erik Velasco Salido en 2015 [23], con el fin de conseguir utilizarlo con eficacia sobre imágenes en color obtenidas con la Kinect v2.

## 3.2. Algoritmo color

### 3.2.1. Diseño

Este algoritmo está constituido por tres partes diferenciadas; una primera parte que consiste en obtener la región de interés del video, otra segunda parte en la que se reproduce de la manera más fiel posible el algoritmo del MIT [27], que consiste, en la amplificación del video para revelar pequeños cambios de color producidos en la cara de una persona a lo largo del tiempo como consecuencia del flujos sanguíneo en la cabeza, y una tercera parte en la cual se obtiene el pulso cardiaco del sujeto mediante el tratamiento del video amplificado en el espacio transformado.

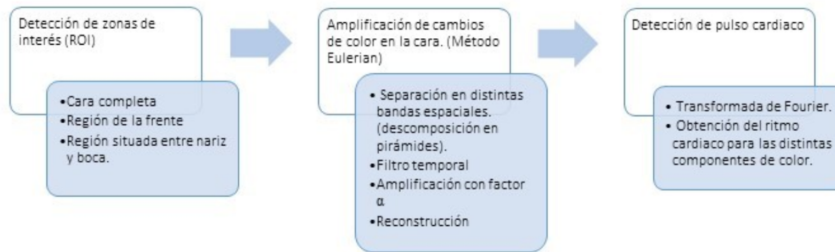


Figura 3.1: Etapas del algoritmo de color

En primer lugar, se obtienen las regiones de interés mediante el algoritmo de Viola-Jones [25]. La elección de unas buenas zonas de interés es primordial en este algoritmo, ya que serán las zonas sobre las que se realice todo el tratamiento, y en función de su estabilidad podremos obtener, o no, unos buenos resultados. Por este motivo, el algoritmo se diseña para trabajar con tres regiones diferentes, con el fin de observar cual de estas la más factible; cara completa, región de la frente, región situada entre nariz y boca. 3.2



Figura 3.2: Zonas de Interés

Una vez obtenida la zona de interés deseada, se realizará la amplificación de cambios de color sobre ella. Para realizar esta amplificación se sigue el método propuesto por el Instituto Tecnológico de Massachusetts 3.3 que consiste en la descomposición

de la secuencia de entrada en diferentes bandas espacio-frecuencias, para posterior aplicación de un filtro temporal en el rango de frecuencias del corazón (40-90 pulsaciones por minuto), para desechar las posibles vibraciones producidas por artefactos en la escena. Las diferentes bandas filtradas se amplifican según un valor de  $\alpha$  variable para hacer perceptibles las variaciones de color en la piel y se reconstruye la señal añadiéndole la original. El vídeo generando nos permitirá observar los cambios de color que surgen debido al flujo sanguíneo, imperceptibles para el ojo humano.

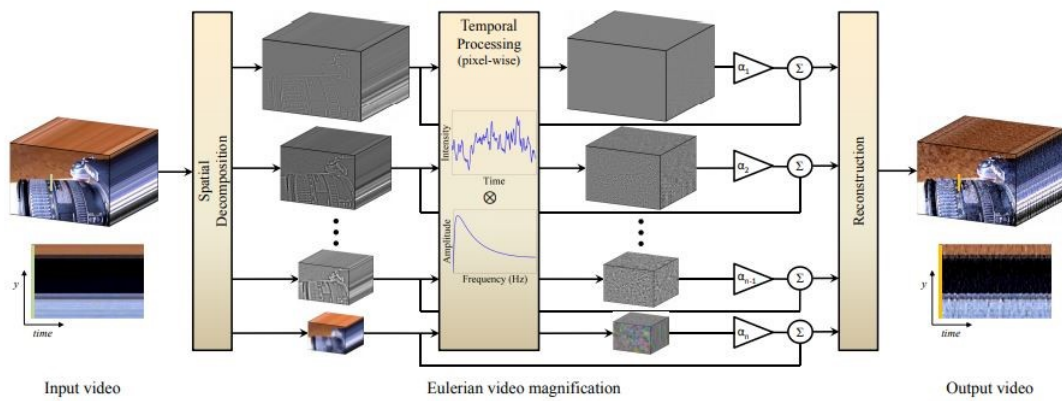


Figura 3.3: Esquema de amplificación de vídeo - Fuente [ref MIT]

Por último, se procede a obtener la frecuencia cardíaca, analizando en el espacio transformado. En este algoritmo se han utilizado dos métodos; mediante la Transformada Discreta de Fourier (DFT), y la Transformada Discreta del coseno (DCT). Sobre cada espacio se calcula la máxima energía para buscar la frecuencia a la que se encuentra dicho máximo. Esta frecuencia será con la cual se calcule el ritmo cardíaco.

### 3.2.2. Desarrollo

El código se ha implementado íntegramente en Matlab 2018. Debido a que se tienen diferentes regiones de interés y distintas formas de tratarlas, se han realizado dos variantes de ejecución de este algoritmo con el fin de comparar resultados, eficacia y velocidades de ejecución y procesamiento.

La primera consiste en recortar las distintas zonas de interés antes del tratamiento y amplificación de los cambios de color, realizando este análisis sobre cada una de las zonas por separado, calculando el pulso cardíaco en cada una de ellas 3.4. La segunda opción o variante, consiste en el tratamiento y amplificación de cambios de color sobre toda la región facial, para, en última instancia, recortar las zonas de interés y sacar

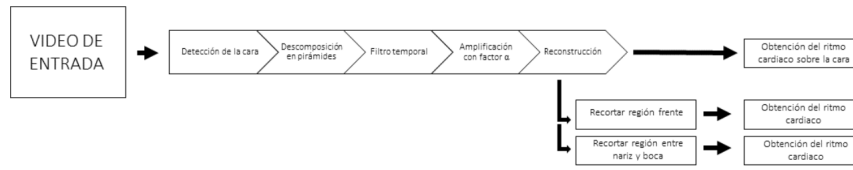


Figura 3.4: Modelo de bloques de la variante 1

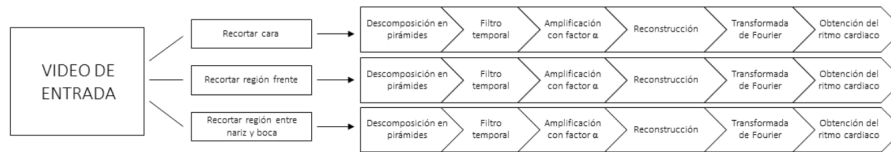


Figura 3.5: Modelo de bloques de la variante 2

sobre ellas el pulso cardíaco [3.5](#).

Una vez esquematizadas ambas variantes y, teniendo en cuenta que ambas contienen los mismos módulos, aunque el tratamiento no se realice en el mismo orden, se procederá a explicar el desarrollo de cada uno de los bloques.

### 3.2.2.1. Detección facial

El primer paso consiste en proceder a la obtención de la zona facial en la primera imagen del vídeo en color. En el caso de vídeos en los cuales la detección no se puede realizar de manera automática, como puede ser el caso de recién nacidos tumbados, se selecciona la cara de forma manual seleccionándola mediante la función IMCROP. Para el resto de casos, la detección facial se realizará mediante el algoritmo de detección de Viola-Jones siguiendo lo expuesto por Balakrishnan [\[25\]](#), se utiliza la función CASCADEOBJECTDETECTOR de Matlab que encuentra la región facial [3.6](#).

Siguiendo la segunda variante, y tomando como región la frente, por ejemplo, se procederá a recortar la región durante todo el vídeo.

Una vez encontrada la región facial, se procede a recortar todas las imágenes que contiene la secuencia del vídeo en la misma zona. Esto funciona ya que se tiene en cuenta que el usuario está colaborando y se mantiene relativamente inmóvil durante toda la grabación. Se ha comprobado que, teniendo esta colaboración por parte del sujeto, no es necesario realizar un seguimiento de la zona facial durante toda la secuencia de vídeo, ya que es más costoso computacionalmente y no hay gran diferencia



Figura 3.6: Cara Detectada

a la hora de obtener el pulso cardíaco con respecto a encontrar la zona facial en la primera imagen y seleccionar dicha zona en el resto de imágenes de la secuencia de vídeo.

Vamos a suponer, para una mejor comprensión del algoritmo, que se sigue el esquema de la segunda variante. Sobre la secuencia de imágenes faciales obtenida, se procede a obtener las dos regiones de interés de la cara 3.2, que se han elegido por su estabilidad. Para ello se ha seguido lo establecido en [2].

Supongamos, entonces, que se elige la región de la frente como región de interés. Se recortará esta región en cada una de las imágenes de la secuencia, obteniendo una secuencia de salida en la cuál solo figura la frente.

#### 3.2.2.2. Descomposición Piramidal

Sobre ella realizaremos la descomposición piramidal para realizar un análisis multiresolución mediante Wavelets. Este tipo de descomposición nos ayuda a reducir el ruido y detectar contornos en ciertas escalas de resolución. Las Wavelets ayudan a realizar un análisis tiempo-frecuencia efectivo. La figura 3.7 nos ayudará a entender mejor el concepto de una descomposición piramidal.

Citando la descripción desarrollada en [6] - La base de la pirámide representa la imagen original, “f” filas y “c” columnas. Cada nivel de la pirámide, al que se accede desde el nivel inmediatamente inferior, es una aproximación de la imagen original, de forma que se cumple que en cada nivel N la imagen aproximación tiene “f/2” filas y “c/2” columnas, es decir, conforme ascendemos de escalón o nivel en la pirámide, las sucesivas aproximaciones a la imagen original van teniendo cada vez menos resolución. Las transformaciones wavelet dan la diferencia entre un escalón y otro de la imagen, es decir, las distintas imágenes de aproximación a partir de la imagen original. En cada uno de estos pasos de resolución espacial respecto a la imagen original, el detalle se separa en tres zonas, caracterizadas por los habitualmente denominados coeficientes

de detalle horizontal, vertical y diagonal. Estos coeficientes de detalle recogen la diferencia de información existente entre dos imágenes aproximación pertenecientes a dos niveles sucesivos de la pirámide - fin de la cita. [6]

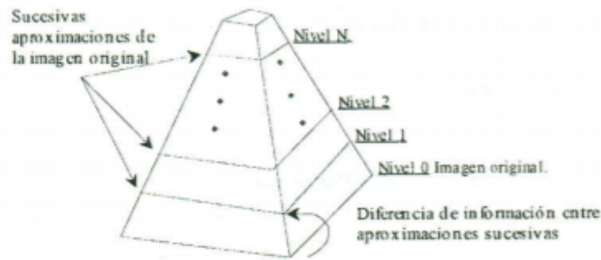


Figura 3.7: Representación piramidal del análisis multiresolución. Fuente [6]

En primera instancia, se debatió la idea de usar Wavelet Gaussianas o Laplacianas y finalmente, debido a que existen estudios con resultados satisfactorios de detección de ritmo cardíaco mediante secuencias de vídeo utilizando una u otra [27], se asumió, que para este análisis el resultado sería similar en ambas. Por este motivo, en el algoritmo una se usa una pirámide Wavelet Gaussiana.

La implementación de esta pirámide se ha realizado con ayuda del toolbox propio de Matlab “matlabPyrTools”. Se realiza una pirámide por cada componente de color (RGB), y para cada imagen de la secuencia. La cantidad de niveles de la pirámide es un parámetro importante en la realización de este algoritmo ya es decisivo en la cantidad de detalle y variación de color de la secuencia de vídeo final. Este parámetro se ha estudiado durante la realización de este trabajo, realizando varios barridos sobre él. Se explicará con más detalle en el apartado de resultados.

El tipo de filtro utilizado es un filtro binomial debido a que el MIT corrobora que es uno de los que mejor funciona [27]. El resultado es la imagen Wavelet Gaussiana, resultante de aplicar la pirámide una determinada serie de veces (niveles de la pirámide gaussiana), por cada imagen de secuencia de vídeo. Esta imagen contiene tres capas, una por cada componente de color.

Esta secuencia de imágenes gaussianas se introduce en una variable 4D, donde las dimensiones son; alto y ancho de la imagen, componentes de color y tiempo (h,w,3,t). De esta manera podremos realizar el siguiente paso del algoritmo, filtrado temporal, de una forma más eficiente.



### 3.2.2.3. Filtrado temporal

El filtrado temporalmente se realiza con un filtro ideal para acotar el rango de frecuencias a las deseadas para la obtención del pulso cardíaco. El ancho del filtro, la frecuencia de corte inferior y la frecuencia de corte superior, variarán dependiendo del video de entrada, debido a que si realizamos un filtrado amplio, de 0.667-1.5Hz, correspondiente a 40 a 90 pulsaciones por minuto, por ejemplo, los resultados tienen mucho más ruido y por tanto mayor variación con respecto a la realidad. Por este motivo, en el siguiente capítulo se hará un barrido sobre los distintos parámetros, con el fin de obtener un código más automático.

### 3.2.2.4. Amplificación y reconstrucción

Una vez filtrada la secuencia de vídeo, se ha implementado una función que realiza la amplificación y reconstrucción. Esta función toma como parámetros de entrada el video sin tratar, el video filtrado y el factor de amplificación. Lo primero que se hace es multiplicar cada capa del video filtrado (RGB), por un parámetro alpha, el cuál iremos variando en el capítulo siguiente para determinar el que nos proporciona unos mejores resultados. Tras este paso, vamos a obtener el vídeo resultante como una suma de cada imagen del video original con su correspondiente imagen tratada y filtrada. En la figura 3.8 se muestran los cambios de color que ahora sí son perceptibles.

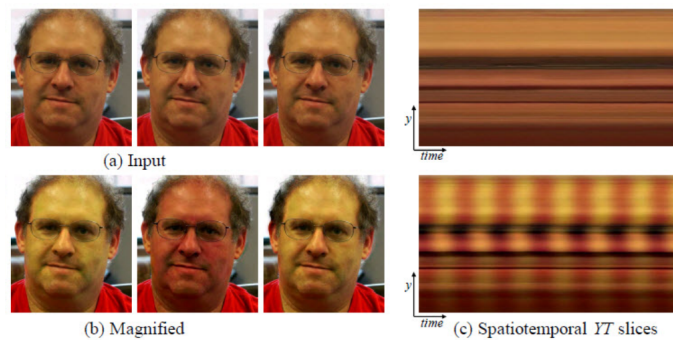


Figura 3.8: Amplificación de variaciones de color

### 3.2.2.5. Cálculo de la frecuencia cardíaca

Para poder calcular el ritmo cardíaco una vez amplificado y reconstruido el vídeo, primero vamos a recortar la sección (en este caso ya la tendríamos recortada por lo que no sería necesario realizar este paso), y se calcularía la media de color, creando un

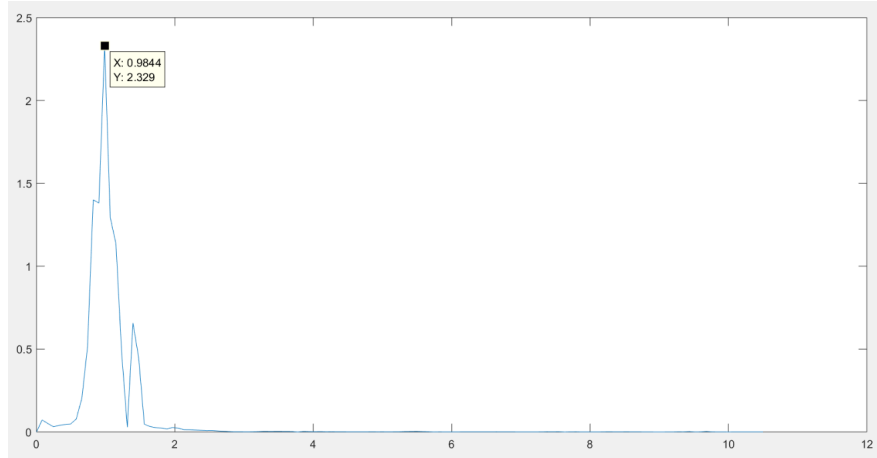


Figura 3.9: Representación de la fft para la extracción del pulso cardíaco

vector, con la media a lo largo del tiempo, por cada capa. Después de esto se realiza la transformada de Fourier para cada uno de los canales, realizamos el módulo y la donde se encuentre la mayor energía será la frecuencia a la que está bombeando la persona, en frecuencia/s, pero tenemos que despreciar la primera componente que nos aparece ya que es la que al realizar la fft se corresponde con la componente continua que no nos interesa. Una vez obtenida la frecuencia de mayor energía, se multiplica este valor por 60 para obtener las pulsaciones por minuto.

Todos estos cálculos se van a realizar para cada zona y para cada componente de color en una función. Cabe destacar que la componente de luminancia (intensidad), revela de una manera muy eficaz las variaciones de color que se producen tras la amplificación.

### 3.3. Algoritmo Luminancia

#### 3.3.1. Diseño

Este algoritmo contiene los mismos bloques y variantes que el algoritmo de color, la única diferencia consiste en convertir del espacio RGB al espacio YCbCr en el primer bloque, detección de la región facial 3.6. De este modo, podremos realizar todo el tratamiento sobre un único canal, consiguiendo una mayor rapidez computacional. La luminancia representa el brillo o las componentes blanca y negra de una señal de vídeo. A partir de la luminancia mediante una media ponderada se obtienen las distintas componentes de color (RGB) de cada pixel, por lo que realizando el tratamiento sobre la luminancia, no estaremos perdiendo información y será más eficiente

que analizar cada canal RGB por separado.

### 3.4. Desarrollos no incluidos

#### 3.4.1. Introducción

En este apartado se resumirán los diferentes desarrollos que se han realizado a lo largo de este trabajo, pero que no han sido incluidos por que o no han tenido resultados del todo satisfactorios o no eran demasiado extensos para poder incluirlos en esta memoria, (este va a ser nuestro caso del algoritmo de movimiento).

Los desarrollos han sido:

- Construcción de un filtro Butterworth para realizar el filtrado temporal. A pesar de que computacionalmente es más rápido, los filtros ideales de banda estrecha obtienen mejores resultados.
- Pirámide Wavelet laplaciana a la hora de descomponer en distintas escalas de resolución.
- PSD como espacio transformado para calcular la frecuencia cardíaca. La Densidad Espectral de Potencia (PSD) da peores resultados que la DCT y FFT.
- Mejora del algoritmo de movimiento [2] para utilizarlo sobre imágenes de color. Se explicará con más detalle en la siguiente subsección.

##### 3.4.1.1. Mejora del algoritmo de movimiento sobre imágenes de color

Debido a que Matlab, que es el entorno donde se ha trabajado, se ha ido actualizando, se ha visto en la necesidad de adaptar el algoritmo de Erik Velasco [23] para Matlab 2018. Además, se ha analizado el movimiento en cada uno de los canales de color, RGB, en contraposición a lo implementado en [23] basado en escala de grises. En el siguiente apartado se analizarán los resultados obtenidos con este algoritmo.

### 3.5. Validación de resultados

#### 3.5.1. Introducción

En esta sección vamos a realizar una primera evaluación de los resultados arrojados por los algoritmos implementados para comprobar la fiabilidad de estos. Para ello, se ha tomado como referencia el algoritmo del MIT [27] y tres de los vídeos que acompañan a dicho algoritmo. 3.10

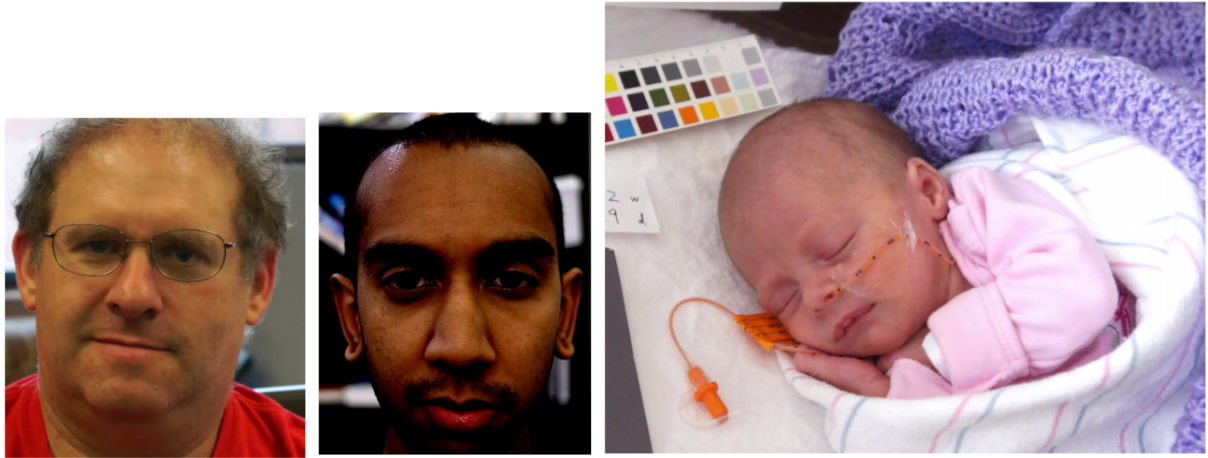


Figura 3.10: Imágenes de prueba

El problema de este código, al igual que del propuesto, es que tiene parámetros variables. Para comprobar resultados usaremos exactamente los que usan en el código del MIT [27].

### 3.5.2. Resultados

Comprobando los resultados obtenidos en 3.1, se puede llegar a observar que se aproximan bastante entre ellos y al Ground-truth. Por este motivo llegamos a la conclusión de que el algoritmo está implementado de manera oportuna.

Vídeo	<i>Ground-truth</i>	Algoritmo MIT + fft	Algoritmo Color + fft	Algoritmo Luminancia + fft	Algoritmo movimiento	Error ( % ) MIT-Color
face_54bpm	54	55.36	54.03	55.36	55	2.402
face_55bpm	55	56.06	56.25	56.06	56.5	-0.339
face_152bpm	152	141	146	141	152	-3.546

Tabla 3.1: Validación de resultados con dataset MIT

### 3.5.3. Conclusiones

Dando pie al siguiente capítulo, deberemos tener en cuenta que si bien es cierto, que los resultados obtenidos en el anterior apartado parecen satisfactorios, estos se han obtenido ejecutando el algoritmo y estableciendo valores fijos a los parámetros sin un criterio sustentado. Por este motivo, el capítulo 4 de evaluación se va a centrar tan solo en realizar un código que contenga el algoritmo de color y el de luminancia. Se realizará sobre todos los parámetros un barrido, con el fin de obtener buenos

resultados. Este barrido tiene la condición fundamental de comprobar la sensibilidad de los distintos parámetros.



## Capítulo 4

# Evaluación

### 4.1. Introducción

Una vez realizado el diseño y desarrollo de los distintos algoritmos y habiendo obtenido, en la última sección del apartado anterior, buenos resultados de la primera valoración de estos con respecto al del Instituto Tecnológico de Massachussets[27], en este capítulo se detallarán los diferentes resultados que se han obtenido tras realizar un barrido completo, sobre los distintos parámetros que componen los algoritmos de color y luminancia, y que serán los factores clave para obtener resultados satisfactorios.

### 4.2. Marco de evaluación

#### 4.2.1. *Ground truth*

El término *Ground truth* se refiere a la precisión de la clasificación del conjunto de entrenamiento para las técnicas de aprendizaje supervisado. Se usa en modelos estadísticos para probar o refutar hipótesis de investigación.

Durante el proceso de grabación las diferentes secuencias de vídeo, se ha medido la frecuencia cardíaca con un pulsometro de muñeca instalado en el smartwatch “Fitbit Versa”, que será nuestro *Ground truth*, para poder comparar las medidas de frecuencia cardíaca dadas por el smartwatch, con las obtenidas por los diferentes algoritmos implementados. Este dispositivo tiene un margen de error despreciable para sujetos en reposo, como lo son los sujetos que se han prestado a este estudio, por lo que tendremos una fuente fiable a la hora de comparar resultados.



Figura 4.1: Smartwatch “Fitbit Versa”

#### 4.2.2. *Dataset*

El *Dataset* que se ha generado está compuesto por dos bloques diferenciados: los vídeos que se han obtenido de la página web del MIT y las secuencias que se han grabado durante la realización de este trabajo. Las secuencias de vídeo grabadas en color tienen una duración de entre 10 y 15 segundos. Estas secuencias se han obtenido mediante varios recursos materiales:

- Webcam integrada en Portatil Spectre 13-af00ns con procesador Inter core i7-8550U, con 8GB de memoria RAM, SSD de 256 GB y sistema operativo de 64 bits. Esta Webcam graba con una resolución de 0.9 mega-píxeles, un tamaño máximo de imagen de 1280x720, y frame rate de 30 frames/s.
- Cámara Kinect v2 con frame rate de 30 frames/s y tamaño de la imagen, para vídeos en color, de 1920x1080.





Figura 4.2: Muestras de los distintos tipos de secuencias que componen el dataset

Como se puede observar en la figura anterior, las secuencias de vídeo que componen el dataset se han realizado a diferentes distancias, posiciones, resoluciones y tipos de piel. En total, tenemos 15 secuencias de vídeo. El nombre del archivo que contiene cada secuencia de imágenes de un vídeo se caracteriza por “NombrePersona\_ppm\_framerate\_tiempo\_webcam/kinect” para que de manera sencilla se tengan localizadas las especificaciones de cada secuencia.

Tabla 4.1: Ground Truth

	Sujetos voluntarios para el estudio														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pulsaciones por minuto	65	78	90	53	68	68	69	74	75	90	60	54	70	75	70

Las secuencias de vídeo de este dataset han sido obtenidas mediante el software en desarrollo de Julia Simón.

### 4.2.3. Métricas

Para la evaluación de los algoritmos se han seguido las siguientes métricas:

- *Ground truth*: Media de las pulsaciones medidas por la *Fitbit Versa* durante la grabación del vídeo.
- *Error*: Es el error entre la el valor del *Ground truth* y la media del pulso arrojado por el algoritmo implementado durante la realización de este trabajo.

$$Error(\%) = \frac{Groundtruth - ValorObtenido}{Groundtruth} \times 100$$

### 4.3. Pruebas y resultados

#### 4.3.1. Pruebas y resultados barrido

El objetivo de este apartado es realizar un barrido exhaustivo sobre todos los parámetros configurables de nuestro algoritmo, con el fin de analizar los resultados y determinar los valores más adecuados para ellos. Los parámetros que se han tenido en cuenta para la realización de este barrido son:

- Alpha: Este es un parámetro del que no se han encontrado grandes referencias o estudios, por ello se va a analizar a partir de que valor de amplificación se hace posible obtener el ritmo cardíaco, y si existe algún valor a partir del cual se estabilicen los resultados
- Bandas de frecuencias: Como se explicó en el capítulo anterior, al realizar el filtrado temporal sobre una banda de frecuencias muy amplia, al tratarse de un filtro ideal, el resultado no siempre es satisfactorio, debido a la influencia de ruido. Por este motivo se van a analizar filtros de banda estrecha a distintas frecuencias y se va a observar el comportamiento de los resultados.
- Regiones de interés antes y después del procesamiento: En este barrido realizaremos el análisis sobre las dos variantes de color explicadas en el capítulo anterior [3.5](#) [3.4](#). Tendremos entonces seis zonas; cara, región de la frente y región entre la nariz y la boca, realizando el procesamiento antes de recortar las zonas, y realizándolo después de recortarlas.
- Diferentes niveles de la pirámide: La cantidad de niveles de la pirámide puede ser crucial a la hora de una buena obtención del ritmo cardíaco. Por este motivo se va a realizar un estudio de este parámetro, comenzando por un nivel  $n=3$  hasta llegar a  $n=6$ , ya que en el estudio realizado por el MIT se utilizan valores en este rango para la detección del pulso cardíaco [\[27\]](#).

Para entender el barrido de manera sencilla se ha representado de manera esquematizada en la siguiente figura, donde la banda de frecuencias se representa como pulsaciones por minuto, para mayor claridad.

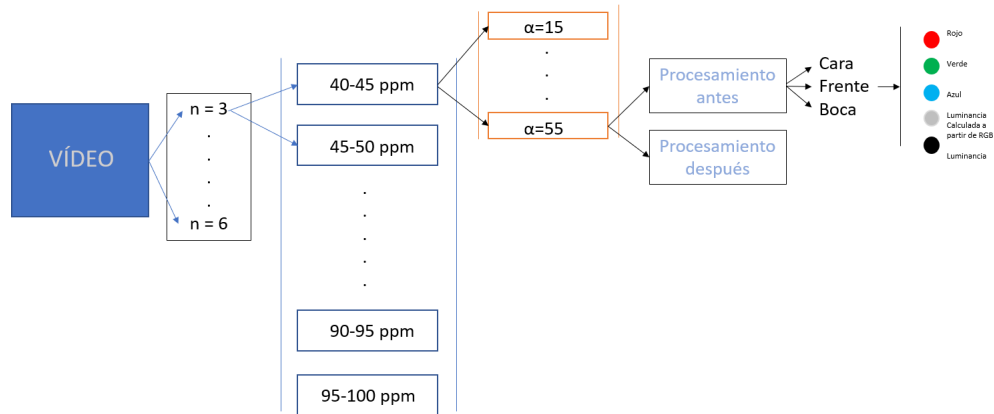


Figura 4.3: Esquema del algoritmo que se llevará a cabo para realizar el barrido

Una vez realizado este barrido, se han obtenido, para cada vídeo, 11 conjuntos de resultados, una por cada banda de frecuencias, de 216x5 resultados (variando los valores de  $\alpha$ , partes de la cara, niveles y canales). Al tratarse de un gran cantidad de resultados, se van a representar en esta memoria tan solo algunos ejemplos de gráficas. Los datos completos están disponibles bajo petición.

Se mostrarán en formato gráfico los resultados para el vídeo de un sujeto con frecuencia cardíaca de 64 pulsaciones por minuto. En concreto se representa el canal rojo para tres bandas de frecuencias distintas (40-45ppm, 60-65ppm y 70-75ppm), con todos los posibles valores de  $\alpha$  y las diferentes zonas de la cara.

Con este análisis, se ha llegado a la conclusión de que la diferencia entre procesar primero las zonas o procesarlas después, es despreciable, por lo que se ha elegido el primer camino, ya que al ser más eficiente nos proporciona mayor rapidez computacional.

Se ha visto que a partir de un cierto valor de  $\alpha$ ,  $\alpha=30$ , se tienen unos resultados bastante estables, y que, en valores por debajo de 5-10 los resultados pueden dar falsos positivos por la poca amplificación por lo que esos valores de  $\alpha$  tan pequeños no se van a utilizar. Por estos motivos, se ha decidido que el algoritmo final tome valores de  $\alpha$  de entre 15 y 30 a intervalos de 5. No se elige un solo valor de  $\alpha$  ya que esto puede provocar falsos positivos.

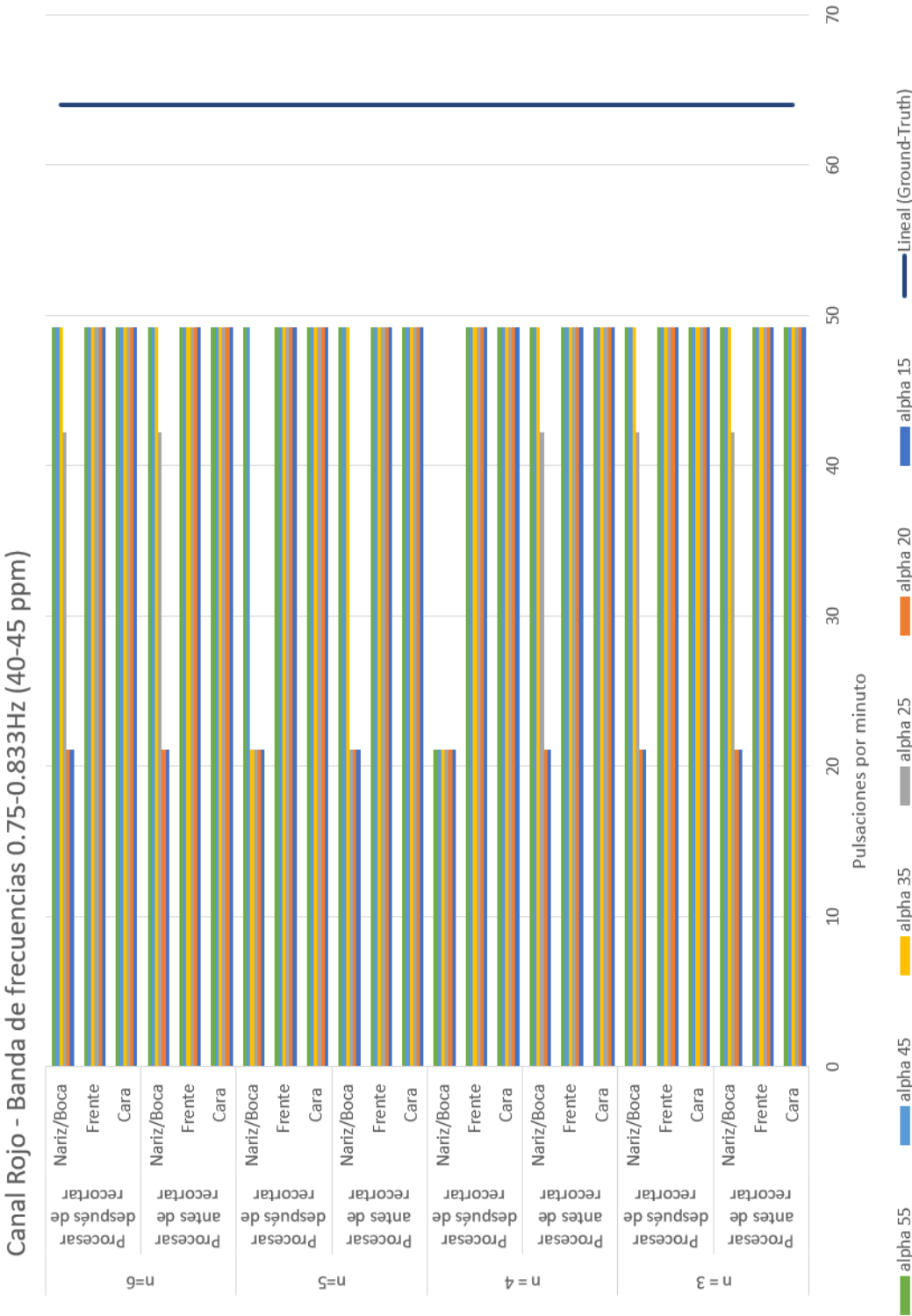


Figura 4.4: Canal rojo - Banda frecuencia 0.75-0.833 Hz



Figura 4.5: Canal rojo - Banda frecuencia 1-1.0833 Hz

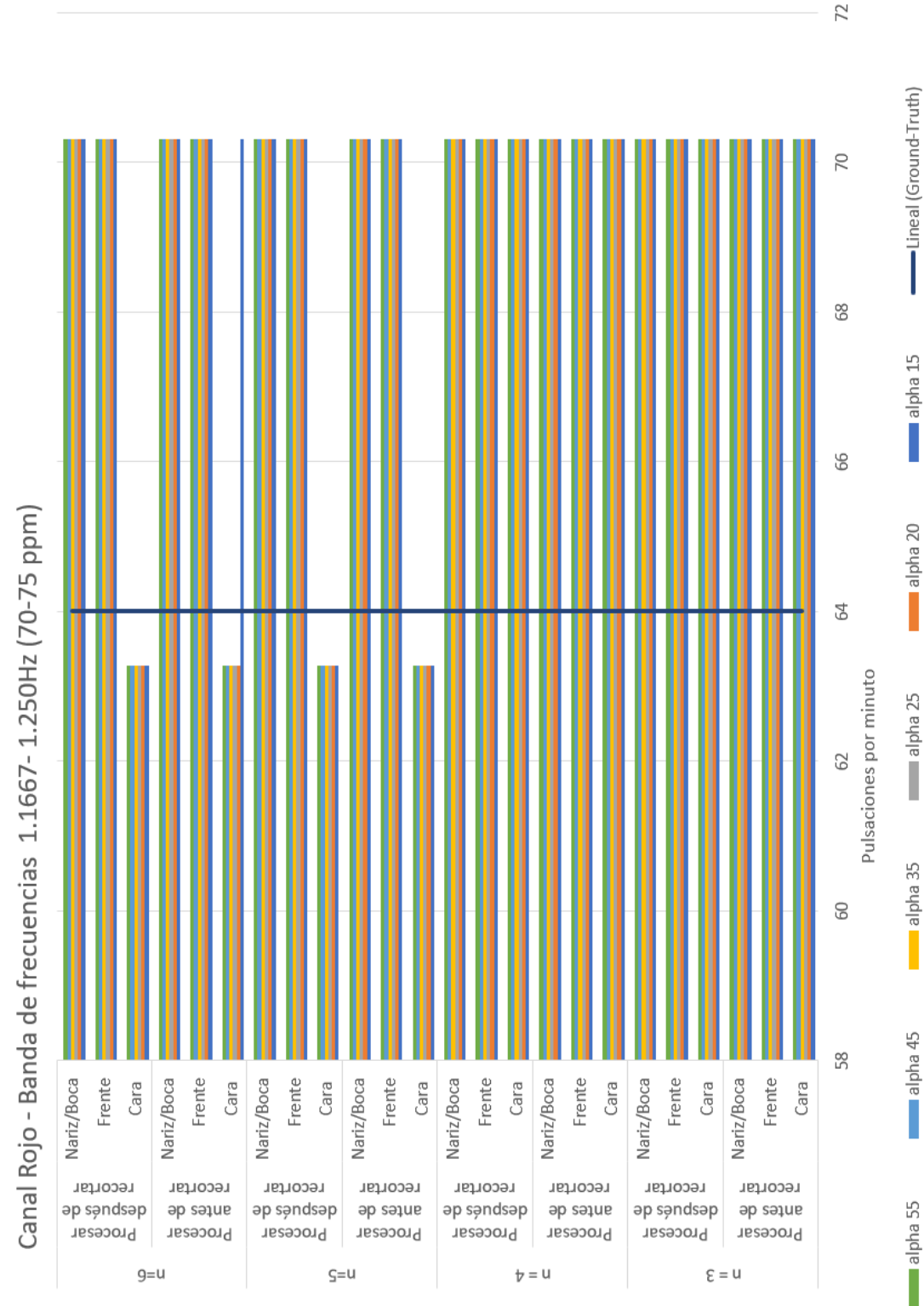


Figura 4.6: Canal rojo - Banda frecuencia 1.1667-1.25 Hz

Aunque el valor de la pirámide, dentro del rango medido, no varía demasiado los resultados, se ha elegido utilizar  $n=4$ , puesto que en los dos vídeos probados es el que más estable se mantiene sin aumentar el coste computacional de forma exagerada.

Parámetro	Valor seleccionado
Nivel de la pirámide	$n=4$
Alpha	15,20,25,30
Procesamiento	Antes de recortar las zonas
Bandas de frecuencia	Banda estrecha (0.0833Hz)

Tabla 4.2: Parámetros Seleccionados

Una vez decididos los parámetros que vamos a usar se ha vuelto a realizar el algoritmo, para obtener nuevos resultados 4.7, para una secuencia de vídeo donde el Ground truth es de 64 pulsaciones por minuto. Para simplificar las tablas, se han visualizado los resultados únicamente para  $\alpha=20$  y  $\alpha=25$ . En las tablas, se podrá ver el valor de las pulsaciones para los canales rojo, verde, azul, luminancia obtenida mediante los canales RGB, y luminancia obtenida por el algoritmo de luminancia implementado, para las distintas regiones de la cara, y para los distintos valores de  $\alpha$ . Tomamos como valores posibles aquellos en los que las pulsaciones obtenidas se encuentran dentro del rango de frecuencias que se está analizando (en 4.7 serían los que están marcados en verde). El valor de pulsaciones real será el que más veces aparezca dentro del rango establecido. Para el caso estudiado, observando las tablas, vemos que el que más veces aparece se encuentra en la banda de frecuencias de 60-65ppm, que es la que se corresponde con el Ground truth, en concreto sería el valor 63,281 que se aproxima mucho al real.

A continuación de manera resumida se explican los pasos a seguir para obtener el valor correcto:

1. Tomamos los posibles valores dentro de cada banda de frecuencias.
2. Realizamos la moda en cada banda de frecuencias y contabilizamos cuantas veces aparece el valor más repetido de cada banda.
3. Seleccionamos el que aparezca un mayor número de veces dentro de la banda adecuada.

Banda de frecuencias 40-45 ppm							
Rojo	Verde	Azul	Y'	Y			
49,21875	49,21875	49,21875	49,21875	7,03125	cara	alpha 20	
49,21875	49,21875	49,21875	49,21875	49,21875		alpha 25	
49,21875	49,21875	49,21875	49,21875	49,21875	frente	alpha 20	
49,21875	49,21875	49,21875	49,21875	49,21875		alpha 25	
21,09375	7,03125	7,03125	11,2359375	7,03125	nariz/boca	alpha 20	
21,09375	49,21875	49,21875	40,809375	7,03125		alpha 25	

Banda de frecuencias 45-50 ppm							
Rojo	Verde	Azul	Y'	Y			
49,21875	56,25	56,25	54,1476563	7,03125	cara	alpha 20	
49,21875	56,25	56,25	54,1476563	49,21875		alpha 25	
42,1875	56,25	56,25	52,0453125	49,21875	frente	alpha 20	
42,1875	56,25	56,25	52,0453125	49,21875		alpha 25	
56,25	7,03125	56,25	27,3585938	7,03125	nariz/boca	alpha 20	
56,25	56,25	56,25	27,3585938	7,03125		alpha 25	

Banda de frecuencias 50-55 ppm							
Rojo	Verde	Azul	Y'	Y			
56,25	49,21875	63,28125	52,9242188	56,25	cara	alpha 20	
56,25	49,21875	63,28125	52,9242188	56,25		alpha 25	
56,25	56,25	63,28125	57,0515625	56,25	frente	alpha 20	
56,25	49,21875	63,28125	52,9242188	56,25		alpha 25	
49,21875	7,03125	49,21875	24,4546875	7,03125	nariz/boca	alpha 20	
49,21875	63,28125	49,21875	57,4734375	7,03125		alpha 25	

Banda de frecuencias 55-60 ppm							
Rojo	Verde	Azul	Y'	Y			
56,25	7,03125	56,25	27,3585938	56,25	cara	alpha 20	
56,25	7,03125	56,25	27,3585938	56,25		alpha 25	
56,25	7,03125	56,25	27,3585938	56,25	frente	alpha 20	
56,25	56,25	56,25	56,25	56,25		alpha 25	
56,25	7,03125	56,25	27,3585938	7,03125	nariz/boca	alpha 20	
56,25	7,03125	56,25	27,3585938	56,25		alpha 25	

Banda de frecuencias 60-65 ppm							
Rojo	Verde	Azul	Y'	Y			
63,28125	7,03125	63,28125	30,2625	63,28125	cara	alpha 20	
63,28125	7,03125	63,28125	30,2625	63,28125		alpha 25	
63,28125	63,28125	63,28125	63,28125	63,28125	frente	alpha 20	
63,28125	63,28125	63,28125	63,28125	63,28125		alpha 25	
63,28125	63,28125	63,28125	63,28125	63,28125	nariz/boca	alpha 20	
63,28125	63,28125	63,28125	63,28125	63,28125		alpha 25	

Banda de frecuencias 65-70 ppm							
Rojo	Verde	Azul	Y'	Y			
70,3125	70,3125	63,28125	69,5109375	70,3125	cara	alpha 20	
70,3125	70,3125	63,28125	69,5109375	70,3125		alpha 25	
70,3125	70,3125	63,28125	69,5109375	63,28125	frente	alpha 20	
70,3125	70,3125	63,28125	69,5109375	63,28125		alpha 25	
70,3125	70,3125	63,28125	69,5109375	70,3125	nariz/boca	alpha 20	
70,3125	70,3125	63,28125	69,5109375	70,3125		alpha 25	

Banda de frecuencias 70-75 ppm							
Rojo	Verde	Azul	Y'	Y			
70,3125	70,3125	63,28125	69,5109375	70,3125	cara	alpha 20	
70,3125	70,3125	63,28125	69,5109375	70,3125		alpha 25	
70,3125	70,3125	63,28125	69,5109375	63,28125	frente	alpha 20	
70,3125	70,3125	63,28125	69,5109375	63,28125		alpha 25	
70,3125	70,3125	63,28125	69,5109375	70,3125	nariz/boca	alpha 20	
70,3125	70,3125	63,28125	69,5109375	70,3125		alpha 25	

Banda de frecuencias 75-80 ppm							
Rojo	Verde	Azul	Y'	Y			
7,03125	77,34375	77,34375	56,3203125	7,03125	cara	alpha 20	
7,03125	77,34375	77,34375	56,3203125	7,03125		alpha 25	
77,34375	77,34375	77,34375	77,34375	77,34375	frente	alpha 20	
77,34375	77,34375	77,34375	77,34375	77,34375		alpha 25	
77,34375	77,34375	7,03125	69,328125	7,03125	nariz/boca	alpha 20	
77,34375	77,34375	77,34375	77,34375	7,03125		alpha 25	

Banda de frecuencias 80-85 ppm							
Rojo	Verde	Azul	Y'	Y			
84,375	84,375	84,375	84,375	7,03125	cara	alpha 20	
84,375	84,375	84,375	84,375	7,03125		alpha 25	
84,375	84,375	84,375	84,375	84,375	frente	alpha 20	
84,375	84,375	84,375	84,375	84,375		alpha 25	
21,09375	84,375	84,375	65,4539063	7,03125	nariz/boca	alpha 20	
21,09375	84,375	84,375	65,4539063	7,03125		alpha 25	

Banda de frecuencias 85-90 ppm							
Rojo	Verde	Azul	Y'	Y			
84,375	84,375	84,375	84,375	7,03125	cara	alpha 20	
84,375	84,375	84,375	84,375	7,03125		alpha 25	
91,40625	84,375	84,375	86,4773438	91,40625	frente	alpha 20	
91,40625	84,375	84,375	86,4773438	91,40625		alpha 25	
21,09375	84,375	84,375	65,4539063	7,03125	nariz/boca	alpha 20	
21,09375	84,375	84,375	65,4539063	7,03125		alpha 25	

Banda de frecuencias 90-95 ppm							
Rojo	Verde	Azul	Y'	Y			
91,40625	98,4375	91,40625	95,5335938	7,03125	cara	alpha 20	
91,40625	98,4375	91,40625	95,5335938	7,03125		alpha 25	
91,40625	98,4375	91,40625	95,5335938	91,40625	frente	alpha 20	
91,40625	98,4375	91,40625	95,5335938	91,40625		alpha 25	
21,09375	91,40625	98,4375	71,184375	7,03125	nariz/boca	alpha 20	
21,09375	91,40625	98,4375	71,184375	7,03125		alpha 25	

Figura 4.7: Gráficas de resultados

Banda de frecuencia (ppm)	Moda (dentro del rango de frecuencia)	Cuantos valores lo cumplen
40-45	40.809	1
45-50	49.219	5
50-55	52.924	3
55-60	56.25	18
60-65	63.281	26
65-70	69.511	6
70-75	70.312	16
75-80	77.344	20
80-85	84.375	22
85-90	86.477	2
90-95	91.406	12

Figura 4.8: Posibles valores de pulsaciones



### 4.3.2. Resultados finales

Teniendo fijos los parámetros tal y como se expone en la tabla 4.2 del apartado anterior, se ha realizado una evaluación del algoritmo usando, además de la Transformada Discreta de Fourier, la Transformada Discreta del Coseno. Hemos añadido secuencias de vídeo de personas con distintos tonos de piel, y grabaciones realizadas con varias resoluciones.

Secuencia	Ground-truth	Resolución video	Obtenido mediante	Pulsaciones con DCT	Pulsaciones con FFT	Tiempo de procesamiento DCT (s)	Tiempo de procesamiento FFT (s)	Error % DCT	Error % FFT
1	75	1280x720	WebCam	72	77.35	19.622	30.883	4	3.133
2	90	1280x720	WebCam	91.20	90.31	20.784	37.501	1.333	1.455
3	54	1280x720	WebCam	54.3	56.25	25.074	38.022	0.555	4.166
4	74	1920x1080	Kinect v2	72.45	74.01	67.881	57.504	2.095	0.013
5	53	1920x1080	Kinect v2	70.3125	75.43	64.152	55.273	32.66	42.32
6	69	1920x1080	Kinect v2	69.8	68.46	59.515	52.489	1.159	0.783

Tabla 4.3: Resultados finales con utilización DCT y FFT

## 4.4. Conclusión

Podemos concluir que los valores son bastante favorables tanto para el caso de la DCT y la FFT. Además los tiempos de procesamiento son de baja latencia, por lo que en un futuro será posible obtenerlo para tiempo real. Además, la DCT es más estable en las bandas de frecuencias en las que se encuentra la frecuencia cardíaca, y menos estable cuando son otras bandas, por tanto, se pueden realizar bandas de barrido de ancho mayor que la FFT, por lo que se considera que otorga mejores resultados.



## Capítulo 5

# Conclusiones y trabajo futuro

### 5.1. Conclusiones

El objetivo principal es implementar un sistema que estime, mediante el análisis de variaciones de color en la piel, la frecuencia cardíaca de un individuo, mediante técnicas no invasivas. Por este motivo la detección de ritmo cardíaco mediante secuencias de vídeo nos otorga una buena solución. La frecuencia cardíaca obtenida mediante los algoritmos implementados durante la realización de este trabajo ha sido, para todos los casos, muy similar a el *Ground truth*, por lo que concluimos que se han implementado de manera satisfactoria. Además, la realización de esta detección sobre secuencias de vídeo en color, nos da la posibilidad de detectar variaciones en la piel de la persona, que pueden ser producidas por problemas arteriales.

### 5.2. Trabajo futuro

A la vista de los resultados que se han obtenido en este trabajo se propone trabajar en:

- Trabajar sobre las componentes independientes (ICA) sobre cada canal para obtener mejores resultados.
- Utilizar la percepción de las variaciones de color para crear un sistema de detección de enfermedades arteriales.
- Conseguir una región de interés que de unos resultados uniformes en todas las personas.
- Trabajar sobre las componentes de color HSI, ya que en él los valores de la piel de las personas entran dentro de un mismo rango, por que puede dar resultados más favorables para personas con piel oscura, o para casos con baja iluminación.



# Bibliografía

- [1] Guha Balakrishnan. *Analyzing pulse from head motions in video*. PhD thesis, Massachusetts Institute of Technology, 2014. 6, 11, 13
- [2] Guha Balakrishnan, Fredo Durand, and John Guttag. Detecting pulse from head motions in video. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3430–3437. IEEE, 2013. 13, 19, 23
- [3] C Barón et al. Balistocardiografo: historia de un instrumento para medir en forma indirecta el desempeño del corazon. *Revista Colombiana de Cardiología*, 16(1):5–10, 2009. 6
- [4] Jeffrey B Bolkhovsky, Christopher G Scully, and Ki H Chon. Statistical analysis of heart rate and heart rate variability monitoring through the use of smart phone cameras. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1610–1613. IEEE, 2012. 11
- [5] David Da He, Eric S Winokur, and Charles G Sodini. A continuous, wearable, and wireless heart monitor using head ballistocardiogram (bcg) and head electrocardiogram (ecg). In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 4729–4732. IEEE, 2011. 13
- [6] M González de Audicana, R García, and A Seco. Fusion de imágenes de diferente resolución espacial utilizando el análisis muuitresolución y las transformaciones wavelet. XIII, 19, 20
- [7] Martin Fischer, Yen Yang Lim, Elaine Lawrence, and Leena K Ganguli. Remotecare: Health monitoring with streaming video. In *Mobile Business, 2008. ICMB'08. 7th International Conference on*, pages 280–286. IEEE, 2008. 9
- [8] Julian P. J. Halcox. Una breve recorrido por la historia del pulso arteria. 5
- [9] William Harvey. *The circulation of the blood*. Cosimo, Inc., 2006. 5

- [10] Bernard L. Horecker. The absorption spectra of hemoglobin and its derivatives in the visible and near infra-red regions by. 2003. 10
- [11] Viktor Kessler, Markus Kächele, Sascha Meudt, Friedhelm Schwenker, and Günther Palm. Machine learning driven heart rate detection with camera photoplethysmography in time domain. In *IAPR Workshop on Artificial Neural Networks in Pattern Recognition*, pages 324–334. Springer, 2016. 11
- [12] Alex Kleiner and Dan Rabinowitz. Video heart rate detection. 11
- [13] Sungjun Kwon, Hyunseok Kim, and Kwang Suk Park. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 2174–2177. IEEE, 2012. 11
- [14] Magdalena Lewandowska, Jacek Rumiński, Tomasz Kocejko, and Jędrzej Nowak. Measuring pulse rate with a webcamâa non-contact method for evaluating cardiac activity. In *Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on*, pages 405–410. IEEE, 2011. 11
- [15] Ioannis Pavlidis and James Levine. Thermal image analysis for polygraph testing. *IEEE Engineering in Medicine and Biology Magazine*, 21(6):56–64, 2002. 9
- [16] Ming-Zher Poh, Daniel McDuff, and Rosalind Picard. A medical mirror for non-contact health monitoring. In *ACM SIGGRAPH 2011 Emerging Technologies*, page 2. ACM, 2011. 11
- [17] Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 18(10):10762–10774, 2010. 11
- [18] Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE transactions on biomedical engineering*, 58(1):7–11, 2011. 11
- [19] T Pursche, J Krajewski, and Reinhard Moeller. Video-based heart rate measurement from human faces. In *Consumer Electronics (ICCE), 2012 IEEE International Conference on*, pages 544–545. IEEE, 2012. 11
- [20] Lorenzo Scalise, Natascia Bernacchia, Ilaria Ercoli, and Paolo Marchionni. Heart rate measurement in neonatal patients using a webcam. In *Medical Measure-*

- ments and Applications Proceedings (MeMeA), 2012 IEEE International Symposium on*, pages 1–4. IEEE, 2012. 11
- [21] Chihiro Takano and Yuji Ohta. Heart rate measurement based on a time-lapse image. *Medical Engineering and Physics*, 29(8):853–857, 2007. 9, 11
- [22] Gill R Tsouri, Survi Kyal, Sohail A Dianat, and Lalilt K Mestha. Constrained independent component analysis approach to nonobtrusive pulse rate measurements. *Journal of biomedical optics*, 17(7):077011, 2012. 11
- [23] Erik Velasco Salido. Detección de ritmo cardiaco mediante vídeo. B.S. thesis, 2015. 13, 15, 23
- [24] Wim Verkrusysse, Lars O Svaasand, and J Stuart Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, 2008. 11
- [25] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2001. 16, 18
- [26] Lan Wei, Yonghong Tian, Yaowei Wang, Touradj Ebrahimi, and Tiejun Huang. Automatic webcam-based human heart rate measurements using laplacian eigenmap. In *Asian Conference on Computer Vision*, pages 281–292. Springer, 2012. 11
- [27] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. Eulerian video magnification for revealing subtle changes in the world. 2012. 9, 11, 16, 20, 23, 24, 27, 30